Proceedings of **ITC-CSCC 2000**, Pusan, Korea

Video Rate Control Using Activity Based Rate Prediction

Hyung-Shin Park, You-Young Jung, Young-Ro Kim, and Sung-Jea Ko
Department of Electronics Engineering, Korea University
5-1 Anam-Dong, Sungbuk-Ku, Seoul 136-701, Korea
Tel: +82-2-3290-3228, Fax: +82-2-928-0179
E-mail: sjko@dali.korea.ac.kr

Abstract: In this paper, an efficient rate control algorithm based on rate prediction is proposed for maintaining a smooth buffer variation and a small buffer size. The proposed method adjusts the quantization scaling factor by using the predicted bit-rate to meet the target bit budget exactly. Experimental result show that the proposed prediction-based rate control scheme can regulate the bit-rate across scene changes more effectively and achieve better PSNR performance than existing rate control mechanisms such as the MPEG-2 Test Model 5 (TM5) and the Adaptive Scene Analysis (ASA).

1. Introduction

For digital video transmitted over a constant bit-rate channel, rate control is important to visual quality. Various video rate control techniques have been proposed to achieve a target bit-rate with consistent visual quality. Recent some techniques use the models derived from previous coded bits for the current macroblock encoding [1]-[8], [10]. Other techniques use models estimated off-line from the training sequence for the current macroblock under a stationary assumption [7]. In aforementioned techniques, the coding parameters and quantizer scaling factors are determined based on information derived from previously coded data. These methods may suffer from degradation at scene changes. To solve this problem, a large buffer can be used to smooth out the variation in compressed data. But, larger buffer size will cause longer frame delay.

In this paper, an efficient rate-prediction method and a video rate control algorithm using the predicted rate are proposed. The proposed method adjusts the quantization stepsize using the bit-rate predicted at each slice or each macroblock in order to distribute bits for uniform quality. One advantage of the approach is that the method is accurate and robust for different sequences even in maintaining a small buffer. Another advantage is that the algorithm can maintain the visual quality of image frames with scene changes uniformly. Simulation results of the proposed technique will be composed to those of TM5 and ASA rate control techniques, which are the most popular rate control techniques.

This paper is organized as follows. In Section 2, the proposed rate prediction and rate control algorithms are presented. Experimental results are given in Section 3. And conclusions are given in Section 4.

2. Proposed Prediction-based Rate Control

Before presenting the proposed rate control scheme, we explain the rate prediction method based on the DCT-based activity.

2.1 Rate Prediction using the DCT-based Activity

In this paper, we use the sum of AC coefficient absolute values as the activity measure for the 8×8 DCT block as follows:

$$A=\sum_{j=1}^{63}\left|C_j\right|, \qquad (1)$$

where C_j is the jth AC coefficient of the 8×8 DCT block. The activity measure in (1) has been known to be the most accurate [3]. The nth order predictor using the activity factor is given by

$$\hat{r}_i = a_0 + a_1 A + a_2 A^2 + \cdots + a_n A^n , \qquad (2)$$

where $\hat{r}_i$ represents the predicted bit-rate for the ith DCT block and $\mathbf{a}=\{a_0, a_1, \cdots, a_n\}$ is the set of the predictor coefficients. To obtain the optimal coefficient set, we use the discrete least square approximation scheme given by

$$E(a_0, a_1, \cdots, a_n) = \sum_{i=1}^{N}\left|r_i - \hat{r}_i\right| , \qquad (3)$$

where $E(a_0, a_1, \cdots, a_n)$ is the error function, N is the number of blocks used for the design of the predictor, and r_i is the real coded bit-rate for the ith block. Minimizing the error function in (3) is equivalent to solving the equation given by

$$0=\frac{\partial}{\partial a_j}\sum_{i=1}^{N}[r_i-(a_0+a_1A+a_2A^2+\cdots+a_nA^n)]^2, \qquad (4)$$
$$j=0,1,\cdots,n.$$

In Section 3, the parameter n is derived experimently using the relationship between the activity factor and the real coded bit-rate. According to the quantization stepsize Q, 31 sets of the predictor coefficients $\{\mathbf{a}_Q, Q=1,2,\cdots,31\}$ are also determined using the test image sequences.

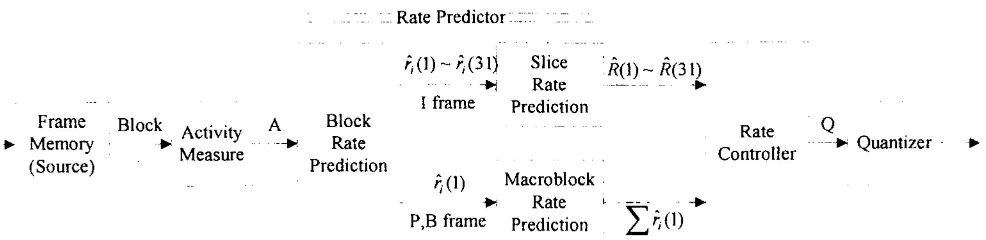


Fig. 1. Functional block diagram of the proposed rate control algorithm.

2.2 Proposed Rate Control Scheme

Since I frame often causes the congestion in the network, we consider separate rate control schemes for I frame and P, B frames.

For I frame, the proposed rate control is performed at each slice. According to quantization stepsize Q, slice bit-rates $\hat{R}(1) \sim \hat{R}(31)$ are first predicted as

$$\hat{R}(Q) = \sum_{i=1}^{M} \hat{r}_i(Q) + H, \quad Q = 1,2,\cdots,31 \tag{5}$$

where M is the number of blocks within a slice, $\hat{r}_i(Q)$ is the predicted block bit-rate associated with the ith block within the slice, and H is the number of coded bits for the slice header. In the second step, the encoding slice bit-rate is adjusted by using the minimum Q satisfying

$$\hat{R}(Q) < T_s, \tag{6}$$

where T_s is the target slice bit-rate given by

$$T_s = \frac{T_I - \sum_{k=1}^{h-h_r} R_k(Q_k)}{h_r}, \tag{7}$$

where T_I represents the target frame bit-rate of I frame which can be determined by TM5, h is total number of slices within a frame, h_r is the number of slices to be processed in the current frame, and $R_k(Q_k)$ is the number of coded bits for the kth slice.

In our algorithm, TM5 is used for rate control for P, B frames. But, the sum of $\hat{r}_i(1)$'s of blocks in the macroblock is utilized instead of the block variance. Fig.1 shows the functional block diagram of the proposed rate control which is consist of the activity measure, the rate predictor, and the rate control algorithm can be added to a classical MPEG coder. It should be noted that it is completely MPEG compatible and transparent to decoders.

3. Simulation and Results

To evaluate the video rate control performance of the proposed algorithm, we compare the results of the proposed technique with those of TM5 and ASA. For our simulation, we use four video sequences, with 110 frames, Flowergarden, Mobile, Tennis, and Football video sequence. These sequences are encoded according to CCIR601 4:2:0 format, 240 lines by 352 pels, and GOP (6,3) coding structure. The picture rate is 25 frames/s and the video data rate is set to 0.6 and 1.2 Mbps.

3.1 Performance of the bit-rate predictor

In order to determine the predictor coefficients, we used aforementioned four images to examine the relationship between activity measures and the number of coded bits. Fig. 2 is experimental results associated with a fixed quantization stepsize Q=4, where each point represents the number of coded bits versus the activity value of the 8×8 MPEG image block. Note that the relationship between the coded bits and the activity value can be well approximated by a piece-wise linear line with five disjoint activity regions, as shown in Fig. 3. In each disjoint region, we derive the piece-wise linear bit-rate predictor given by

$$\hat{r}_i = a_0 + a_1 A \tag{8}$$

This method to design the predictor is similar to the method proposed by Cheng and Hang [4].

Table 1 shows predictor coefficients associated with Q=4. Fig. 4(a) shows the relationship between the real coded bits and the predicted bits obtained by using the prediction coefficients in Table1. For comparison, the real coded bits versus the variance values of TM5 is shown in Fig. 4(b). It is seen that the proposed method produces the linear approximation, while TM5 using variance does not.

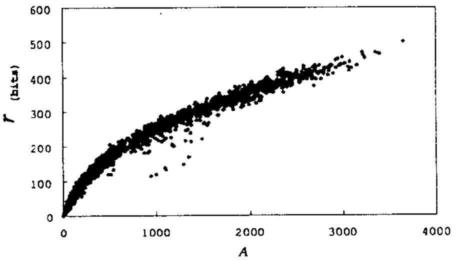
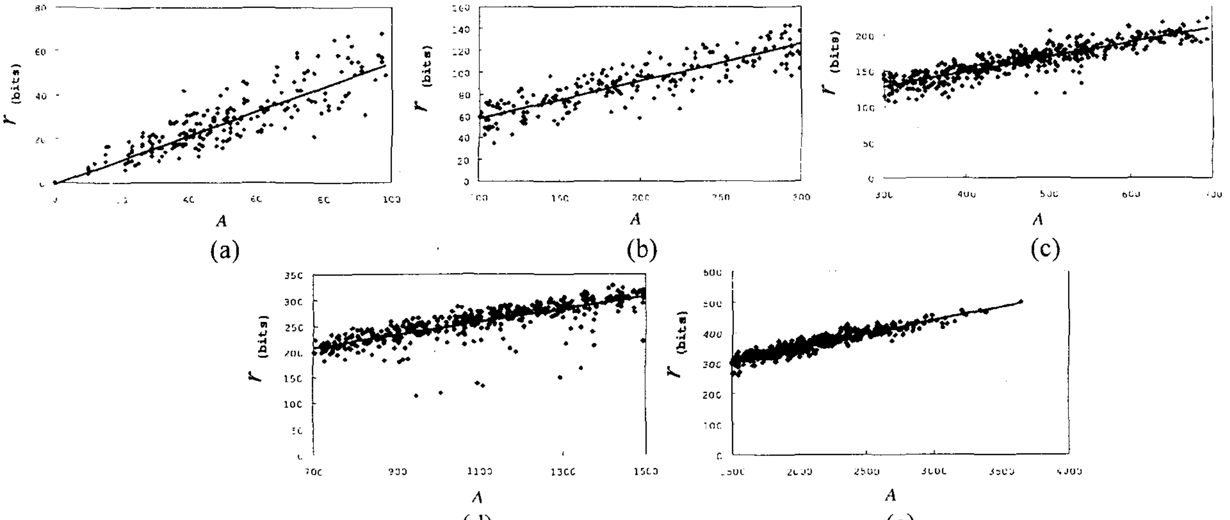


Fig. 2. Encoded bits versus activity for Q=4.

Table 1. Prediction coefficients associated with each activity region.

Activity region A	Prediction coefficients	
	a_0	a_1
0 ~ 99	0.4879	0.5514
100 ~ 299	24.40	0.3375
300 ~ 699	65.89	0.2075
700 ~ 1499	120.41	0.1246
1500 ~	176.68	0.0882



(a) (b) (c)

(d) (e)

Fig. 3. Rate prediction model with Q=4.
(a) A=0~99. (b) A=100~299. (c) A=300~699. (d) A=700~1499. (e) A=1500 ~.

Thus, the piece-wise bit-rate predictors are used for the proposed rate control algorithm.

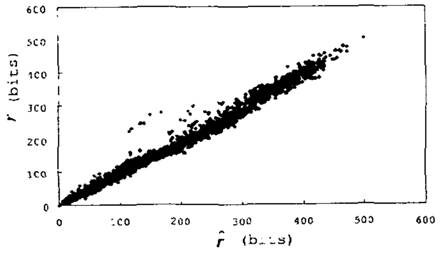
3.2 Simulation of the proposed rate control algorithm

The peak signal to noise ratio (PSNR) and the frame bit counts associated with Flowergarden, Mix video sequences are presented in Fig. 5 and Fig. 6. Here, Mix is a video sequence which is composed of the Flowergarden and the Mobile sequences at the video data rate 0.6 Mbps. Thus, Mix has a few scene changes (in frames 21, 41, 61, and 81), although Flowergarden has no scene change. Fig. 5(b) shows that the bit-rate generated by TM5 exceeds the desired I frame target bit-rate at I frames between the 59th and the 95th. In Fig. 6(b), the encoded bit-rates of I frame obtained by TM5 and ASA are also over the I frame target bit-rate at several I frames. On the other hand, the proposed scheme controls the output bits rather precisely to match the designed target.

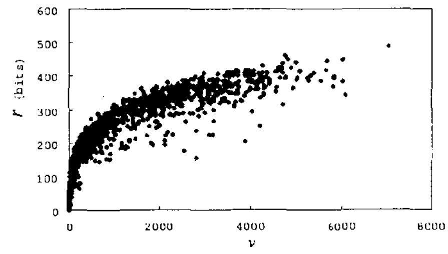
We calculated the average deviation for actual numbers of bits of I frames from the target T_I [2], [6] which is given by

$$D = \frac{1}{N}\sum_{k=1}^{N}\frac{|T_I - T_k|}{T_I}\times 100\ (\%) \qquad (9)$$

where N is the number of I frames in the tested sequence, T_I is the target bit-rate of I frame, and T_k is the real coded bits of the kth I frame.



(a)



(b)

Fig. 4. Coded bits versus activity. (a) Coded bits versus predicted bits. (b) Coded bits versus variance.

Table 2. Average deviation from the target bit-rate of I frame (in percent). (R: the video data rate)

Image Sequence	D(%) (R=0.6 Mbps)			D(%) (R=1.2 Mbps)		
	TM5	ASA	Proposed	TM5	ASA	Proposed
Flowergarden	10.25	2.54	0.96	10.08	9.89	0.24
Mobile	12.26	8.03	1.41	8.05	8.95	0.66
Tennis	11.18	18.60	3.85	19.46	25.61	2.62
Football	22.78	11.41	1.40	30.83	18.26	0.82
Mix	10.53	3.72	1.56	13.08	14.72	0.44

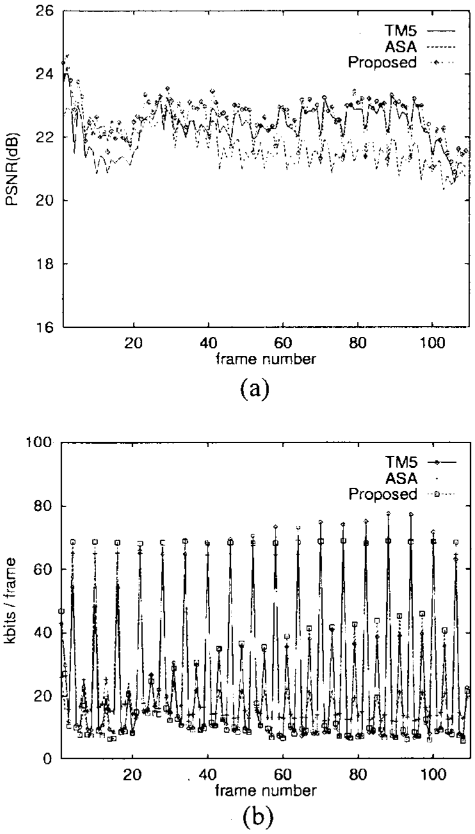
Table 3. Average PSNR.

Image Sequence	PSNR(dB) (R=0.6 Mbps)			PSNR(dB) (R=1.2 Mbps)		
	TM5	ASA	Proposed	TM5	ASA	Proposed
Flowergarden	22.32	21.75	22.60	25.92	26.47	26.57
Mobile	20.93	20.74	21.50	24.08	26.02	25.01
Tennis	28.92	30.04	29.75	32.32	33.95	33.19
Football	25.27	26.57	25.91	28.34	30.19	29.20
Mix	21.52	21.61	22.00	24.79	25.99	28.21

Table 2 illustrates that the proposed rate control scheme can regulate the bit-rate more effectively than TM5 and ASA. The proposed scheme also achieves better PSNR performance than TM5 and ASA as shown in Table 3. The proposed one produces smaller coded bits than TM5 and ASA especially under scene changes. The results in this section indicate that the proposed one can not only avoid the buffer overflow, but also produce uniform visual quality.

4. Conclusion

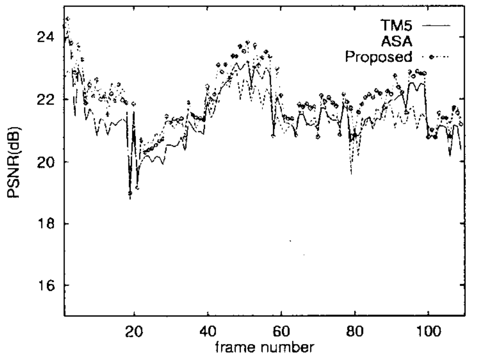
In this paper, we have proposed a video rate control algorithm using activity based rate prediction. It was shown that the proposed algorithm exhibits better performance than existing rate control schemes such as TM5 and ASA.



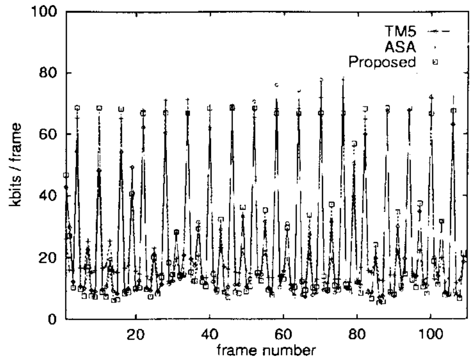
(a)

(b)

Fig. 5. Results of the flowergarden video sequence at the video data rate 0.6 Mbps. (a) PSNR (dB). (b) Bit-rate.



(a)



(b)

Fig. 6. Results of the Mix video sequence at the video data rate 0.6 Mbps. (a) PSNR (dB). (b) Bit-rate.

References

[1] M. R. Pickering and J. F. Arnold, "A perceptually efficient VBR rate control algorithm," *IEEE Trans. Image Proc.*, vol. 3, no. 5, pp. 527-532, Sep. 1994.

[2] ISO/MPEG II, "Test model 5," Doc. AVC-491, Apr. 1990.

[3] W. -Y. Sun, H. -M. Hang, and C. -B. Fong, "Scene adaptive parameters selection for MPEG syntax based HDTV coding," in *Int'l Workshop on HDTV'93*, Ottawa, Canada, Oct. 1993.

[4] J. -B. Cheng and H. -M. Hang, "Adaptive piecewise linear bits estimation model for MPEG based video coding," *Journal of Visual Communication and Image Representation*, vol. 8, no. 1, pp. 51-67, Mar. 1997.

[5] D. J. Reininger and D. Raychaudhuri, "Bandwidth renegotiation for VBR video over ATM networks," *IEEE J. Select. Areas Commun.*, vol. 14, no. 6, pp. 1076-1085, Aug. 1996.

[6] W. Ding and B. Liu, "Rate control of MPEG video coding and recoding by rate-quantization modeling," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 6, no. 1, pp. 12-20, Feb. 1996.

[7] A. Puri and R. Aravind, "Motion-compensated video coding with adaptive perceptual quantization," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 1, no 4, pp. 351-361, Dec. 1991.

[8] M. Hamdi, J. W. Roberts, and P. Rolin, "Rate control for VBR video coders in broad-band networks," *IEEE J. Select. Areas Commun.*, vol. 15, no. 6, pp. 1040-1051, Aug. 1997.

[9] S. B. Gelfand and C. S. Ravishankar, "Tree-structured piece-wise linear adaptive filter," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1907-1922, 1993.

[10] K. -C. Fan and K. -S. Kan, "An active scene analysis-based approach for pseudoconstant bit-rate video coding," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 8, no. 2, pp. 159-170, Apr. 1998.